Gulf of Mexico Red Snapper
Consensus Summary Report

Prepared by the SEDAR 7 Review Panel for the Gulf of Mexico Fishery Management Council

Edited by P.L. Cordue for
SEDAR 7, 4-7 April 2005
New Orleans, LA

Executive summary

The SEDAR 7 Review Workshop met in New Orleans, LA from April 4–7, 2005 to review the stock assessment of red snapper in the Gulf of Mexico. The first day consisted primarily of presentations by the Assessment Team covering the Data Workshop, the two Assessment Workshops, and their preferred base case assessment. During the second and third days, the workshop reviewed the assessment by addressing the terms of reference for the Review Workshop, including the consideration of additional model runs. On the final day, preliminary drafts of the Consensus Summary Report and the Advisory Report were reviewed.

The SEDAR for red snapper has extended over more than 12 months, during which time the Assessment Team and other Data Workshop and Assessment Workshop participants have worked towards producing a credible and reliable stock assessment. The red snapper assessment has been more challenging than the original participants could have envisaged. There were many challenges: being able to fully understand and duplicate the methods and data used in 1999 assessment; exploring alternative stock hypotheses and eventually moving from a single stock model to a two stock model; collating and analyzing the many relevant data sources to provide indices appropriate to single stock and two stock models; constructing a catch history (for multiple fisheries, including discard estimates) extending to the "dawn" of the fishery (1872); undertaking numerous assessment runs using four different stock assessment methods; and choosing a base case assessment to further develop and present to the Review Workshop.

The Review Panel was impressed by the quantity and quality of the work which had gone into the red snapper stock assessment. The presentations of the Assessment Team on the first day were well structured and clear. The information provided, through the presentations, and in response to questions, gave an excellent basis for the Panel's subsequent deliberations and collaboration with the Assessment Team.

Two changes to the base case assessment were made during the Review Workshop. These were suggested by the Panel and agreed to by the Assessment Team. First, age-0 snapper were reintroduced into the model. The Panel understood the argument in support of excluding this age class in that density dependent compensation could extend to even higher ages. However, in the scientific judgment of the Panel, it is not prudent to assume that density dependent compensation can completely override the mortality induced by the shrimp fishery on age-0 red snapper.

The second change was to include higher recruitment scenarios in the projections of the base case. Recruitment estimates over the last 20 years are highly variable, but on average are above the level predicted by the stock-recruitment relationship. Three alternative recruitment scenarios were recommended for projections, using either: the spawner-recruitment relationship; recent average recruitment (last 20 years); or an even higher average recruitment level (obtained from a sensitivity run). In terms of predicting short-term future recruitment levels, the Panel preferred, on the balance of probabilities, the use of average estimated recruitment over the last 20 years (with benchmarks recalculated to be consistent with that level).

The Advisory Report was finalized after the Review Workshop by the Assessment Team. Runs without age-0 snapper are included in that report together with the Review Workshop's base case. The Assessment Team included the runs to honor the Assessment Workshop agreement. The Review Panel believe that these runs are useful to illustrate the sensitivity of the assessment results to the exclusion of age-0 snapper but should not be used for the baseline assessment from which management advice is derived.

1. Introduction

1.1 Time and Place

The SEDAR 7 Review Workshop (RW) met in New Orleans, LA from April 4–7, 2005.

1.2 Terms of Reference for the Review Workshop

1. Evaluate the adequacy and appropriateness of all data used in the assessment and state whether or not the data are scientifically sound;

2. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population parameters such as abundance, biomass, and exploitation and state whether or not the methods are scientifically sound;

3. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population benchmarks (*e.g., MSY, Fmsy, Bmsy, MSST, MFMT, or their proxies*) and required management parameters (*e.g., ABC*) and state whether or not the methods are scientifically sound;

4. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status and, if appropriate, evaluate stock rebuilding; state whether or not the methods are scientifically sound;

5. Ensure that required assessment results (*as listed in the SEDAR Stock Assessment Report Outline*) are clearly and accurately presented in the Stock Assessment Report and that such results are consistent with the Review Panel's decisions regarding adequacy, appropriateness, and application of the data and methods;

6. Evaluate the performance of the Data and Assessment Workshops with regard to their respective Terms of Reference, and state whether or not the Terms of Reference for those previous workshops are adequately addressed in the Stock Assessment Report;

7. Review data and assessment workshop research and monitoring recommendations and make any additional recommendations warranted;

8. Prepare a Peer Review Consensus Report summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. (Drafted by the Panel during the Review Workshop with a final report due three weeks after the workshop ends.)

1.3 List of Participants

Panelists	Affiliation
Cordue, Patrick (Chair)	CIE
Babcock, Elizabeth	NGO; Univ. Miami
Blanchet, Harry	GMFMC SAP; LA DFW
Murphy, Mike	GMFMC SAP; FL FWCC
Nugent, Mike	GMFMC AP
Pilling, Graham	CIE
Prager, Michael	NMFS/SEFSC Beaufort
Rose, Kenneth	GMFMC SAP; LSU
Sissenwine, Michael	NOAA Fisheries

Presenters	
Brooks, Liz	NMFS/SEFSC Miami
Cass-Calay, Shannon	NMFS/SEFSC Miami
Nichols, Scott	NMFS/Pascagoula Lab
Porch, Clay	NMFS/SEFSC Miami
Scott, Jerry	SEFSC
Turner, Steve	NMFS/SEFSC Miami

Council Members	Affiliation
Minton, Vernon	GMFMC/Alabama DNR
Walker, Bobbi	GMFMC

Staff	Affiliation
John Carmichael	SEDAR Coordinator
Dawn Aring	GMFMC
Steven Atran	GMFMC
Loyd Darby	SEFSC

Observers	Affiliation
Bailey, Danica	Louisiana Dept. of Wildlife & Fisheries
Crabtree, Roy	NMFS/SERO
Cufone, Marianne	Gulf Restoration Network
Hano, Brett	Louisiana Dept. of Wildlife & Fisheries
Merriner, John	NMFS/Beaufort Lab
Nance, Jim	NMFS/Galveston
Powers, Joseph	NMFS/SEFSC Miami
Steele, Phil	NMFS/SERO
Thompson, Nancy	NMFS/SEFSC Miami
Viles, Aaron	Gulf Restoration Network
Zales II, Bob	Panama City Boatmen’s Association

1.4 Review Workshop working papers

Document Number	Document Title	Authors
SEDAR7-RW 1	Application of the age-structured assessment model CATCHEM to the U.S. Gulf of Mexico red snapper fishery since 1962.	Porch, C. E.
SEDAR7-RW 2	Revised assessments of Gulf of Mexico red snapper during 1984-2003 using a gulf-wide implementation of ASAP	Cass-Calay, S. L. and G. A. Diaz.
SEDAR7-RW 3	Revised assessments of Gulf of Mexico red snapper during 1962-2003 using a gulf-wide implementation of an age-structured assessment program (ASAP).	Cass-Calay, S. L., G. A. Diaz, and J. S. Nowlis.
SEDAR7-RW 4	Assessments of red snapper stocks in the eastern and western Gulf of Mexico using an age structured assessment procedure (ASAP)	Ortiz, M. and S. L. Cass-Calay.
SEDAR7-RW 5	Revised bootstrapping of a gulf-wide implementation of an age-structured assessment procedure (ASAP) for red snapper (*Lutjanus campechanus*) from 1962 to 2003.	Nowlis, J. S. and S. L. Cass Calay.
SEDAR7-RW 6	An age-structured stock reduction analysis (SRA) model for Gulf of Mexico red snapper that accounts for uncertainty over the ages of density-dependent natural mortality.	McAllister, M. K.
SEDAR7-RW 7	Alternate fishery-independent larval indices of abundance for red snapper.	Hanisko, D. S., J. Lyczkowski-Shultz, and W. Ingram.
SEDAR7-RW 8	Alternative estimates of the yield of red snapper from the Gulf of Mexico recreational fishery.	Turner, S. C.

2. Terms of Reference

2.1 Background

The RW is usually the third meeting in the SEDAR process. However, for red snapper, the Data Workshop (DW) was followed by *two* Assessment Workshops (AW). The SEDAR for red snapper has extended over more than 12 months, during which time the Assessment Team and other DW and AW participants have worked towards producing a credible and reliable stock assessment. The red snapper assessment has been more challenging than the original participants could have envisaged.

There were many challenges: being able to fully understand and duplicate methods and data used in the 1999 assessment; exploring alternative stock hypotheses and eventually moving from a single stock model to a two stock model; collating and analyzing the many relevant data sources to provide indices appropriate to single stock and two stock models; constructing a catch history (for multiple fisheries, including discard estimates) extending to the "dawn" of the fishery (1872); undertaking numerous assessment runs using four different stock assessment methods; choosing a base case assessment to further develop and present to the Review Workshop.

The Panel was impressed by the quantity and quality of the work which had gone into the red snapper stock assessment. The presentations of the Assessment Team on the RW's first day were well structured and clear. The information provided, through the presentations, and in response to questions, gave an excellent basis for the Panel's subsequent deliberations and collaboration with the Assessment Team.

During the RW some small deficiencies were noted by the Panel in the proposed base case assessment. The Assessment Team were willing and able, during the RW, to make the minor changes to the base case necessary to address the Panel's concerns. The changes were minor in terms of implementation, although potentially quite important for projections and evaluation of management options. The changes to the assessment are discussed in the following section under the appropriate terms of reference.

2.2 Review of the Panel's deliberations

This section addresses, in order, each of the eight Terms of Reference for the RW (see Section 1.2).

1. Evaluate the adequacy and appropriateness of all data used in the assessment and state whether or not the data are scientifically sound;

The RW's overall conclusion was that the SEDAR process had thoroughly considered the full range of potential sources of data. The flexibility of the AW's preferred assessment method, CATCHEM, allows the unusually complex and diverse array of available data to be assimilated within the assessment model. The RW did not identify inappropriate use of data (i.e., in this sense the data are scientifically sound), except with regard to the issue of choice of the youngest age within the model and its justification (see the discussion on the stock-recruitment relationship below).

The red snapper assessment uses information on (1) distribution and stock structure, (2) growth and reproduction, (3) natural mortality, (4) stock and recruitment relationship, (5) fishery landings and bycatch/discards, (6) age composition of catches and bycatch, and (7) indices of abundance. There is a complex and varied array of data available to address these information categories, but in most cases the available data are incomplete (e.g., in terms of temporal or spatial coverage) such that it is necessary to impute some missing data, and innovative approaches are needed to derive information. In some cases, such derived information is commonly observed for other fisheries, which is preferable (e.g., observations of discards including samples of age composition).

Distribution and stock structure: There are major fishing grounds in the eastern Gulf of Mexico (GoM), western GoM, and Campeche Bank off Mexico. At present, the US fishery is excluded from the fishing ground off Mexico, although historically, it was a major source of US landings. There is sufficient inferential information (e.g., genetics, otolith microchemistry) to support treating these as separate stocks, although the degree of reproductive isolation and mixing of fishes originating from the three areas is unknown.

Growth and reproduction: Since the late 1990s, there has been a tremendous increase in the number of age determinations of red snappers. These data provided a strong basis for estimating a new growth function, which was done taking account of the potential biasing effect of minimum size regulations. However, the relatively short period of time over which a large number of aging samples were collected means that for most years in the assessment, ages needed to be inferred, thus introducing uncertainty. Relatively few of the age samples collected for 2003 were available for inclusion in the assessment.

Another concern is that age sampling, in some years, may not be sufficiently representative of the catch. Also, a portion of the age and length sampling, in some years, has been taken on an opportunistic basis, rather than as part of a program to broadly and representatively sample the overall harvest. Effects of changes in sampling regimes as well as the ability of opportunistic sampling to characterize the size or age of harvest in a fishery have not been closely examined in this process.

Since the previous assessment, new data have been produced on the fecundity of red snapper, although older fish are poorly represented. Little difference between the eastern and western GoM was detected. A single fecundity at age function was fitted and used in the assessment.

Natural mortality: The assessment used instantaneous natural mortality rates of 1.0, 0.6 and 0.1 for ages 0, 1, and 2 years old and older (2+). The 2+ estimate is based on the longevity of the species (over 50 years) and has not changed in the current assessment. The DW reviewed new analyses on the mortality rate of age-0 and age-1 fish. While none of these analyses were conclusive, the DW agreed that the evidence was sufficient to use natural mortality rates for age-0 and age-1 fish (as given above) that are double the rates used in the previous assessment. While the RW accepted the rationale for increasing the estimates of age-0 and age-1 natural mortality, it noted that these changes were important as they lessen the impact of bycatch of young fish relative to the impact of directed fishing.

Stock-recruitment relationship: As is almost always the case, the stock-recruitment (S-R) relationship is empirically estimated by fitting to derived estimates of spawning stock size (S) and recruitment (R). In the case of the red snapper assessment, the fitting is done within the CATCHEM model, with several assumptions (consistent with previous assessments) to constrain the fit. However, it is necessary to specify the age at which recruitment occurs. The AW specified the age at recruitment as age 1 (compared to age 0 for previous assessments). The AW report (page 13, Methods) states that this approach "essentially assumes that the bycatch mortality rate is negligible compared to mortality rate owing to natural density-dependent processes during the first year of life." The DW did not identify data that was relevant to this assumption, nor does the AW report justify it. However, the RW was informed that there is evidence that the period of density dependent compensation extends through age 0 and possibly age 1. Thus treating mortality during age 0, but not age 1, as part of the compensatory recruitment process was considered a compromise. However, the RW included the age-0 bycatch in the base case due to the factors discussed under RW-TOR 2 below.

Fishery landings and bycatch/discards: Commercial landings and recreational catches have been reasonably well documented by systematic data collection programs since 1963 and 1981, respectively. In response to a recommendation of the DW, sporadic sources of commercial landings data were used to construct a catch history beginning in 1872 when the fishery is presumed to have begun. Recreational catches prior to 1981 were inferred by assuming that catches were proportional to human population in coastal areas, estimated from census data from 1900. In the assessment model, recreational catches were assumed to begin in 1946. While estimates prior to systematic data collection programs are particularly uncertain, the RW accepted them as being plausible and useful, and it did not suggest any alternatives.

It is noteworthy that shrimp fishing effort data is usually not available at the relatively precise depth and location scales necessary for direct bycatch estimation, which means that effort must be estimated using catch per unit effort data from interviewed fishing trips. The recent decline in the number of interviews and differences of the spatial distribution of the fishing trips those interviews cover over time is a concern. Also, it appears that estimates of fishing effort and fishing power are confounded such that it is difficult to estimate trends in the latter, although they have almost certainly occurred.

Unfortunately, there is relatively little data on discards, such as from scientific observers aboard fishing vessels. Closed season logbook data in 2001-2002 was deemed to provide some useful information on discards, although the quality of data from "self reporting" is difficult to judge. Data on recreational discards is routinely collected by interviewing "intercepted" anglers (also a form of self reporting). In general, discard estimates for commercial finfish and shrimp fisheries had to be derived from a relatively sparse set of data on discards based on assumptions that are difficult to verify. However, the RW accepted the estimates as a necessary and appropriate use of the data in order to take account of discards in the assessment.

Age composition of catches and bycatch: Age composition data have been collected sporadically, with a large number of samples collected during 1998-2002. The available data were used in the assessment. Fortunately, CATCHEM is flexible enough to not require complete age composition data. Thus, age compositions are derived within the model constrained by data when and where it is available. The RW accepted this approach and expressed concern that the extent, representativeness, and efficiency of the current sampling design should be examined.

Indices of abundance: Three fishery independent (larval survey, trawl survey, video survey) and two fishery dependent (recreational fishery, commercial longline fishery) indices of abundance are currently available. An additional longline survey was available for only a limited time period, so was not included in the base model. Separate indices were constructed for the western and eastern GoM. Trawl survey data for the eastern GoM is limited because of "hard bottom" that is not suitable for trawling. Thus, alternative fishery independent indices are desirable for the east. Fishery dependent indices were standardized using a commonly used General Linear Modeling framework. The RW agreed that the indices were appropriate for use in the assessment. However, it noted that the true uncertainty in the relationship between the larval index and spawning stock size is likely to be larger than is captured by the sampling coefficient of variation.

General comment about data collection: The RW noted that there are many sources of useful data, and that recent enhancements to data collection programs are encouraging. Unfortunately, it also noted that relatively short term data collection efforts (e.g., it appears that enhancements begun in the late 1990s may be dissipating) are less valuable than long term systematic commitments to building the time series that are the backbone of assessments. Fortunately, a flexible modeling framework is available for the red snapper assessment. In effect, missing data is imputed within the assessment model. However, data based on direct observations are more reliable. The lack of observer data on discards is a particular concern.

2. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population parameters such as abundance, biomass, and exploitation and state whether or not the methods are scientifically sound;

The Panel generally endorsed the methods used in the assessment and considered them to be scientifically sound. The one important exception was the decision of the AW to omit age-0 red snapper from the assessment model (see shrimp bycatch discussion below). The Panel was impressed by the large number of runs which had been performed and the use of multiple assessment methods.

The AW considered results from four different assessment methods: ASAP, SRA, VPA, and CATCHEM. The SRA and VPA models were primarily used in exploratory analysis. ASAP was used in previous red snapper assessments and the original intention of the AW was to update the assessment using a modified version of ASAP. Modifications to ASAP were needed to accommodate new data, and in particular the “ultra-historical” catch series (i.e., starting in 1872). Unfortunately, ASAP exhibited instability when it used the ultra-historical catch series and to a lesser extent the shorter time series (1962-2003 and 1984-2003). Further modifications to ASAP, reduced, but did not eliminate the instability. The AW chose CATCHEM to provide the base case stock assessment.

CATCHEM is in many ways a generalization of the ASAP approach, with more flexibility, better mathematical rigor due to internalizing the catch-at-age fitting, and the ability to model geographic substructure. In particular, it can deal with multiple time series with limited spatial and temporal coverage. Parameter estimates are obtained from a modified maximum likelihood best fit to the data. When fully developed, it is anticipated that CATCHEM will be able to provide fully Bayesian stock assessments for red snapper (with interval estimates obtained from marginal posterior distributions). However, the current assessment provided only point estimates (from the mode of the joint posterior distribution).

The AW report contained relatively few diagnostics and several of the Panel’s requests to the Assessment Team related to the provision of further diagnostics (see Section 3 and Appendix A).

Two stock model: The AW chose to adopt a two-stock model with separate eastern and western stocks. No mixing is assumed between the stocks. This is a model assumption that may be violated, but there is little data currently available for estimating mixing rates.

Goodness-of-fit: Fishery landings were closely matched by the model, an expected feature due to the low CVs of most of these data sets. In general, the model provided good fits to the fishery-dependent and fishery-independent abundance indices, although the shrimp by-catch was not fitted well in early years when CVs were high; and larval indices were generally poorly fitted.

Stock-recruitment relation: The RW shared the concern of the AW over the reliability of the estimated relationship between spawners and recruits, given that estimates of recruitment are highest when the stock is thought to be most depleted. The RW speculated that the stock recruitment function could be quite different today than it was 100 years ago.

Capture (fishing) rates: The estimated age composition of the catch was highly truncated in all but the longline fishery samples. The assessment model attributed much of this to strongly peaked selectivities in all but the longline fishery that displayed a logistic selectivity pattern. The RW investigated estimates of age- and year-specific fishing rates for each of the fisheries to inspect if they were at realistic levels to explain the age composition data. Discussions about the age-composition and fishing rates included thoughts about whether older fish were historically found in near shore waters and were vulnerable to the fishery during the ultra-historic period or whether there is a natural ontogenetic movement of fish to deeper water as they age.

Shrimp by-catch of age-0 fish: The base case recommended by the AW did not include age-0 red snapper. The RW examined the effect of including these fish in the analysis. The logic behind the decision to include or exclude these from the analysis is based on beliefs about the timing and strength of density dependent effects on survival. The RW was unable to comment on the age at which compensatory recruitment processes are complete. However, even if there were data that provided sound evidence that compensation occurs throughout age 0, it would be inappropriate to conclude that bycatch mortality of age-0 fish is insignificant. Doing so, not only requires that compensatory recruitment processes extend through age 0, but also that these processes assert such strong control that the fit of the S-R function would be expected to be extremely tight. Clearly, this is not true for red snapper (probably not for any species). Furthermore, it assumes that S is in the asymptotic region of the S-R function where density dependent compensation is strong, not at low levels of S where compensation is weak. The RW also noted that it is not aware of any other assessment where the possibility that density dependent compensatory processes occurring simultaneously with density independent mortality from fishing (either discards or retained catch) was considered justification for treating the mortality from fishing as insignificant, nor is there a reason to think that the red snapper situation is unique. The RW concluded that the base case model should include age-0 snapper. The RW recommends that future assessments model post-recruitment density dependent mortality, as this is critical for determining the impact of shrimp trawl bycatch on red snapper rebuilding.

3. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population benchmarks (e.g., MSY, Fmsy, Bmsy, MSST, MFMT, or their proxies) and required management parameters (e.g., ABC) and state whether or not the methods are scientifically sound;

The RW agreed that the methods used to derive population benchmarks by the AW were appropriate and scientifically sound. The RW endorsed the AW's view that the actual benchmarks are an emergent property of the harvest strategy; the value of MSY is conditional on selectivity patterns of the gears used in the fishery. Choices about selectivity and benchmark construction lead to some of the biggest differences in statements on stock status in the results. As a result, it is necessary to state clearly what the selected benchmark values are conditional upon.

MSY and SPR benchmarks for the RW base case are provided in the Advisory Report. The RW noted that the particular population benchmarks to be applied are policy dependent. The strategies defined by the Gulf of Mexico Fishery Management Council as possible or practical, and how the Council allocates quota among competing user groups, will define the final benchmarks to be calculated for assessment.

The RW concurred with the conclusion of the AW that, due to uncertainties over the true underlying stock-recruitment function and the underlying patterns in the fishery, spawner per recruit (SPR) benchmark levels may be more robust to these uncertainties. 30%SPR, which has already been employed by the Council, is relatively insensitive to benchmarks derived from a stock-recruitment function. Note, SPR benchmarks are consistent with MSY concepts as estimates of both F_{MSY} and B_{MSY} can be inferred from an SPR.

There is a need to test whether selected or alternative benchmarks are robust to sources of uncertainty within the process. The use of management strategy evaluation would be useful to identify alternative robust red snapper population benchmarks. See recommendations for future work.

4. *Evaluate the adequacy, appropriateness, and application of the methods used to project future population status and, if appropriate, evaluate stock rebuilding; state whether or not the methods are scientifically sound;*

The AW report presented a number of projections from the CATCHEM base case. Only deterministic projections were presented, calculated at the mode of the joint posterior distribution, and using the estimated S-R function to predict future recruitments. These projections were done with several scenarios about the amount of future effort in the shrimp trawl fishery: (1) no shrimp, implying that there would be no snapper bycatch in the shrimp trawl fishery; (2) linked, implying that the effort in the directed and bycatch fisheries would remain at their current proportions; and (3) current shrimp, implying that the effort in the shrimp trawl fishery and the closed season handline fishery would remain at current levels, while the directed fisheries would be reduced proportionally. For isopleth calculations, additional levels of shrimp bycatch reduction were considered, including a 40% reduction in the shrimp effort.

The methods used to project population status and evaluate rebuilding were adequate, appropriate and scientifically sound, and were presented clearly. Ideally, the projections should be stochastic, so that it is possible to estimate probabilities of rebuilding and other performance indicators, but the RW recognizes that the stochastic and deterministic projections generally result in similar management advice. The AW did not calculate an Acceptable Biological Catch (ABC) for each stock as required by their terms of reference. However, the RW considers that it will not be possible to calculate the ABCs without clear guidance from the Council on the level of shrimp trawl bycatch that should be assumed in the calculations. In evaluating rebuilding, the AW should also have recalculated the mean generation time with the new biological information available since the last assessment.

The RW considered that the greatest source of uncertainty in the projections was the assumption that was made about future recruitments. The assessment estimated recruitments that were higher than the estimated pristine recruitment (R_0) in recent years when the spawning stock biomass was very low. Thus, using the estimated S-R function to predict future recruitments implies that future recruitments will be lower than the recruitments seen in the last two decades. To address this uncertainty, the Panel requested that projections be done with three different assumed stock recruitment relationships: (1) R_0 predicted from the base case model fit; (2) R_0 set equal to the average of recruitments from 1984–2003; (3) R_0 set equal to the value estimated in the sensitivity analysis run which began in 1984. In each case, steepness should be kept at the value estimated in the base case. Scenarios 2 and 3 are intended to address the possibility that the recent high recruitments were caused by a long-term shift toward higher productivity of red snapper. To be consistent, the benchmarks were calculated based on the assumed S-R function in each scenario. The Panel considered that the scenario based on recent average recruitments was most likely, and should be considered the base case for the projections. However, the RW was not confident that the actual stock recruit dynamics are well represented by any of the scenarios. Therefore, these projections should only be considered plausible in the short time frame (5 to 10 years). The three scenarios should provide reasonable bounds on the uncertainty about future recruitments.

5. Ensure that required assessment results (as listed in the SEDAR Stock Assessment Report Outline) are clearly and accurately presented in the Stock Assessment Report and that such results are consistent with the Review Panel's decisions regarding adequacy, appropriateness, and application of the data and methods;

The RW commends the AW for a clear and well-written report that concisely reflected a very complicated set of analyses and a complex set of deliberations.

The AW report generally followed the suggested report outline. The RW noted that the report was well-written and was mostly clear in what decisions were made and the rationales for these decisions. The AW report was concise, considering the complexity of the assessment workshop deliberations, and clearly cited the supporting documents at appropriate places. The report appeared to be well-balanced.

The Panel had several minor editorial comments about the Stock Assessment Report. These minor comments included: the need for more detailed discussion related to the use of SPR rather than biomass-based benchmarks, more information on why age-0 red snapper by-catch was not explicitly included, a simple statement of recommended ABC, and a clear explanation of how (effective) spawning biomass was computed. The RW also noted that the research recommendations were scattered among various sections, with only those related to shrimp by-catch explicitly noted in the table of contents.

The RW noted that the wording of RW-TOR 5 was somewhat confusing. Above, we have addressed whether the AW report followed the SEDAR Stock Assessment Report Outline, and "clearly and accurately presented" stock assessment results. The remainder of TOR 5, requires that we address if "such results are consistent with the Review Panel's decisions regarding adequacy, appropriateness, and application of the data and methods". However, the assessment results presented in the AW report pre-date the decisions of the RW. We suggest that TOR 5 be reviewed (see Section 5).

6. Evaluate the performance of the Data and Assessment Workshops with regard to their respective Terms of Reference, and state whether or not the Terms of Reference for those previous workshops are adequately addressed in the Stock Assessment Report;

Data Workshop Terms of Reference

The DW participants compiled and reviewed a truly voluminous amount of information on red snapper biology and fisheries in the GoM. Their report discussed in detail the appropriate terms of reference. They discussed the quality and reliability of the available data, considered alternative assessment methods, and provided research recommendations. The details of the workshop process overall were well-documented in the DW Report. Given the diverse sources and forms of information available and the time allotted, the consensus of the RW was that the DW report adequately addressed its Terms of Reference.

Each DW term of reference is considered below.

1. Evaluate stock structure and develop a unit stock definition.

The life history sub-group of the DW supported development of a two-stock model for GoM red snapper, supported by evidence from several sources. That discussion is well-documented in the DW report.

2. Evaluate the quality and reliability of life-history information (age, growth, natural mortality, reproductive characteristics); develop models to describe growth, maturation, and fecundity by age, sex, or length as appropriate.

This was generally well done. The RW noted that the derivation of the ultra-historical time series was motivated by trying to obtain better contrast to estimate steepness in the S-R function. Assuming constant life history parameters over such an extended timeframe is of course problematic, as is estimating natural mortality for ages 0 and 1.

3. Evaluate the quality and reliability of fishery-independent measures of abundance; develop indices of population abundance by appropriate strata (e.g., age, size, and fishery) for use in assessment modeling.

The RW noted that the indices of abundance seemed to treat variability only as sample variability, not considering other sources of variation. The RW mentioned the changing geographic range of the SEAMAP survey as one potential source – it is preferable that surveys cover the entire range of the species, so that variability across the range can be captured, as well as density values in specific portions of the range.

4. Evaluate the quality and reliability of fishery-dependent measures of abundance; develop indices of population abundance by appropriate strata for use in assessment modeling.

The DW did not directly address issues regarding changes in catchability due to technological improvement in recreational or commercial fisheries over time, which the RW considered to be a potentially significant factor. Sensitivity runs during the RW were not able to resolve the importance of this factor, but it is taken to be a subject for future research.

5. Evaluate the quality and reliability of commercial and recreational fishery-dependent data for determining harvest and discard by species; develop estimates of total annual catch including both landings and discard removals.

The DW participants spent considerable energy constructing a reasonable long term history of catch in the fishery to help assess the status of the current stock against the unfished condition. The RW noted that fishery-dependent data on discards is problematic, since it is self-reported except in the cases of direct observer data. Self-reporting can be biased in either direction, and for many causes.

6. Evaluate the quality and reliability of data available for characterizing the size and age distribution of the catch (landings and discard); characterize commercial and recreational landings and discards by size and age.

This was done. The available data from recreational and commercial fisheries were compiled for the assessment.

7. Evaluate the quality and reliability of available data for estimating the impacts of management actions.

This term of reference was not directly addressed in the DW report and it was not entirely clear to the RW how it should have been. Clearly, management actions can affect data and its interpretation (e.g., changes in minimum sizes affect interpretation of size and age frequencies and must be accounted for in an assessment model), but it is less clear what type of data, by itself, can be used to estimate the effects of management actions. Certainly, an assessment model, using whatever data are available, can be used to investigate whether management actions have had a measurable impact on a stock. However, in order to be able to measure the effects of a management action, it is necessary for the action to have greater effect than the noise in the signal. In the presence of noisy data, that may require management actions with a larger impact than have been seen in the past.

The effect of management actions on the shrimp fishery, with the introduction of BRDs, was noted as one data set relevant to this term of reference. The DW had identified this and the model used fleets with and without BRDs.

8. Recommend assessment methods and models that are appropriate given the quality and scope of the data sets reviewed and management requirements.

This was adequately, though not explicitly, covered in the DW report. The RW noted that the DW and AWs were part of a "continuum" – data and methods are inextricably linked.

9. Provide recommendations for future research (research, sampling, monitoring, and assessment).

Three recommendations were noted in the DW report by the RW.

10. Prepare complete documentation of workshop actions and decisions, and generate a data workshop report (Section II. of the SEDAR assessment report).

The RW noted that the DW report was adequate in respect of this rather ambitious term of reference ("complete" is never really achieved).

Assessment Workshop

The AW participants collectively spent more effort and utilized more information in the current analysis than has been done for any prior red snapper assessment, and possibly for any assessment of any kind in the GoM. The details of the workshop process overall were well-documented in the AW Report and the excellent accompanying Proceedings document. Several alternative models were developed and reviewed, with reports available in the AW documents. Given the diverse sources and forms of information available and the time allotted, the consensus of the RW was that the AW report more than adequately addressed the Terms of Reference. A few minor editorial changes were suggested for the report, but that was not seen as a significant factor.

Each AW term of reference is considered below.

> *1. Select several appropriate modeling approaches based on: 1) available data sources, 2) parameters and values required to manage the stock, and 3) recommendations of the Data Workshop – especially including consideration of possible eastern and western stock units; develop and solve population models incorporating the most recent scientifically sound data.*

The RW was impressed with the range of methods considered and employed during the AWs.

> *2. Select a preferred model approach that will be used to provide estimates of population parameters and stock status; provide complete justification for the selected model as well as a review of those methods pursued but ultimately rejected as a preferred approach.*

The selected model is based on well-developed theory and was reviewed by several independent assessment experts during the two AWs and the RW. The RW noted that CATCHEM, the method used in base case assessments, was not fully developed. It preferred that development of the model be more complete (e.g., including standard diagnostics and stochastic projection options) before being used in an actual stock assessment. However, the RW accepted that in the current case and circumstances it was necessary to use CATCHEM since alternative models had unacceptable limitations, such as being unstable when the entire history of the fishery was included in the analysis.

Past assessments concluded that the status of the stock had varied little in the recent time period, the period from which the majority of information was available for input into stock assessments. The AW participants evaluated several methods to construct a reasonable assessment of the status of the current stock against the unfished condition. The limited information available for the early period of these fisheries was found to create problems for the ASAP model used in prior assessments. Another approach, stock reduction analysis (SRA) was attempted, but was not recommended for use when making management recommendations. Rather, it was seen as a useful exploration tool for testing alternative assumptions. The preferred model, CATCHEM, is a more generalized form of ASAP, with greater ability to include information from multiple sources and to include different fleets fishing at different rates on different segments of the stock.

The AW preferred this model because, among other properties, it was able to reasonably model the fishery over the entire history without additional ad hoc inputs.

3. Provide measures of model performance, reliability, and goodness of fit.

Standardized residuals were not provided in the AW report, but some information was provided at the RW. The multinomial assumption for catch at age data appeared to be violated. This should be addressed in future assessments. Reliability of model output needs to be based on reliability of estimation procedures, and how usable it is for providing estimates of future stock structure. (Simulation studies could improve measures, but some sensitivities have been run during the RW to measure stability relative to some input assumptions.) In the view of the Panel the existing AW report was somewhat deficient with regard to diagnostics.

4. Estimate values for and provide tables of relevant stock parameters (abundance, biomass, fishery selectivity, stock-recruitment relationship, etc; include values by age and year where appropriate).

This was adequately done in the AW report.

5. Consider sources of uncertainty related to input data, modeling approach, and model configuration. Provide appropriate and representative measures of precision for stock parameter estimates.

The first part of this term of reference was well addressed by the AWs. However, interval estimates were not provided for parameters. The current implementation of CATCHEM is somewhat inefficient and given the current speed of available computers it is not possible, in a reasonable timeframe, to provide marginal posterior distributions (and hence creditability intervals) for parameters.

6. Prepare sensitivity runs or consider other modeling approaches to examine the reliability of input data sources.

This term of reference perhaps needs to be reworded to clarify how model runs can test the “reliability” as opposed to the “consistency” of input data. However, the RW comment with regard to the AW performance on this term of reference follows.

It is important to understand which indices and other data have greatest influence on the outputs. If precision is poor, but outputs are strongly influenced by that input, then there should be reservations about those outputs. Sensitivity runs to examine the robustness and reliability of the estimates with respect to the input data sets are important. A paragraph or two that stated why the data provided the results they did would have been useful in the AW report.

7. Provide Yield-per-Recruit and Stock-Recruitment analyses.

This was done.

8. Provide complete SFA criteria: evaluate existing SFA benchmarks, estimate values for alternative SFA benchmarks if appropriate, and estimate SFA benchmarks (MSY, Fmsy, Bmsy, MSST, and MFMT) if not previously estimated; develop stock control rules.

The AW report went to some pains to describe implications of different selectivities on SFA benchmarks, ABC, and future stock conditions. There were three sets of SFA benchmarks provided. While the information on the implications of these benchmarks was provided in the report in the form of isopleth diagrams, the RW felt that tabular formulation of a subset of benchmarks would also be useful. The rebuilding schedule is dependent on policy decision on appropriate reference points. Policy decisions make important differences in terms of distribution of TAC. If projections had been developed, they would need to have been done for each possible policy selection, which was thought to be beyond the purview of the AW. Full development of rebuilding plans would have been inefficient prior to selection of appropriate reference points by the Council, and is easily and swiftly done after that selection.

9. Provide declarations of stock status relative to SFA benchmarks: MSY, Fmsy, Bmsy, MSST, MFMT (or their proxies if appropriate).

This was adequately addressed.

10. Estimate the Allowable Biological Catch (ABC) for each stock if appropriate.

This was addressed. No singular value was estimated, but several acceptable catch scenarios were presented, including an infinite number in isopleth diagrams.

*11. Estimate probable future stock conditions and develop rebuilding schedules if warranted; include estimates of generation time. Calculate rebuilding analyses under the following future exploitation possibilities: F=0, F=current, F=current*0.25, F=current*0.5, F=current*0.75.*

In the AW base case, future recruitment was modeled deterministically at the level of the estimated S-R function. Rebuilding plans were not explicitly examined, but were implicit in the isopleth diagrams for the many scenarios which were evaluated. Rebuilding schedules are dependent on policy decision and their associated reference points. Full development of a rebuilding plan would have been inefficient prior to a policy decision, but is easily done after that selection. Mean generation time was not re-calculated (but will be).

The Panel was satisfied that this term of reference was adequately addressed, but did request that projections be done at two alternative higher levels of mean recruitment (and evaluated relative to benchmarks consistent with the higher recruitment levels).

12. Evaluate the impacts of current management actions, with emphasis on determining progress toward stated management goals.

Current policies were included as one of the many scenarios evaluated in the AW report.

13. Provide recommendations for future research and data collection (field and assessment); be specific if possible in describing sampling design and recommended sampling intensity.

The RW noted various recommendations in the AW report.

14. Provide thorough justification for any deviations from recommendations of the Data Workshop or subsequent modification of data sources provided by the Data Workshop.

Deviations were adequately documented.

15. Fully and completely document all activities in writing:
Draft Section III of the SEDAR Stock Assessment Report;
Provide required tables of estimated values;
Prepare a first draft of the Advisory Report based on the Assessment Workshop's recommended base assessment run for consideration by the Review Panel

All reports and documentation were fully accomplished, except that development of the first draft of the Advisory Report was continued during the RW. This was to accommodate extra work performed after the second AW and during the RW.

7. Review data and assessment workshop research and monitoring recommendations and make any additional recommendations warranted;

The RW reviewed recommendations of the DW and AW, and has also made its own recommendations for research that could improve future assessments. The RW joins the AW in emphasizing that it is critical that suitable planning be done before large-scale research programs are conducted. Initial planning workshops and simulation studies can ensure that subsequent research will contribute the information most needed to resolve important questions in red snapper assessment. The more complex or expensive the proposed research, the more important this recommendation becomes.

Some of the following research recommendations are marked [D] or [A] or both. The symbol indicates that all or part of the corresponding recommendation was adapted from recommendations of the SEDAR 7 Data Workshop or Assessment Workshop.

1. *Data on shrimp fishery*. The RW recognized the importance of obtaining better estimates of fishing effort in the shrimp fishery, which might be done through vessel monitoring systems, electronic logbooks, or otherwise [A]. Also, the RW recommends that the statistical design and extent of the shrimp-trawl observer program be reviewed to ensure that the bycatch data collected are appropriate and sufficient for stock assessment.

2. *Independent estimates of mortality rates*. Direct estimation of mortality rates through tagging would reduce uncertainty in future assessments [A].

3. *Fishing power*. Research is recommended to estimate (independently of any stock assessment) changes in catchability q by gear over time. The RW believes that the introduction of GPS and marine chart-plotting equipment is likely to have increased fishing power substantially for some modes of fishing. Independent collection of data on fishing effort would provide valuable data for assessment and relieve the need to estimate catchability changes.

4. *Stock structure*. Research (e.g., tagging, otolith analysis) is recommended to better describe stock structure and mixing rates. Research should include a review of oceanographic data to see whether transport from the Campeche Banks could reasonably be supplying important numbers of larvae to the western Gulf stock [A].

5. *Spawning-stock index*. Given the many factors that can mask relationship of larvae to spawners, the value of the larval indices should be reviewed.

6. *Spatial distribution at age*. The RW recommends study of the age structure observed from longlines (survey and fishery), to clarify geographic distribution of fish as they age.

7. *Density dependence*. Research could clarify the magnitude and timing of density dependent compensation in juveniles by estimating survival (from age-0 to age-1 year) at different densities of juvenile abundance [A].

8. *Ecosystem concerns*. The RW recommends that the management objectives for the fishery complex (shrimp, red snapper, vermilion snapper, etc.) be formalized. Simulation studies could usefully identify and evaluate appropriate management strategies (including use of various reference points) and corresponding assessment modeling approaches. Research could also test the hypothesis that red snapper production is enhanced in some way by increased shrimp trawling [A].

9. *Assessment modeling*. The RW's recommendations for assessment modeling are made while recognizing that technology is currently limiting (the power of current small computers is marginal for the given model complexity). (a) Future assessments should include interval estimates on parameters and status indicators. (b) More diagnostic and output information should be provided in future assessment reports (e.g., plots or tables of F at age and plots of standardized residuals). (c) Extensive simulation tests of assessment models are recommended to examine accuracy, precision, and robustness [A].

10. *Age sampling*. The RW recommends that representative sampling of age- and length-composition of red snapper be conducted consistently across area, time, and gear.

11. *Fecundity at age*. The RW noted that few fecundity samples were available from older fish, and recommends that more such samples be collected.

12. *Model implementation*. The RW recommends that the assessment model's recruitment submodel be generalized to allow various options on the timing of bycatch mortality relative to density dependent compensation (see AW-8).

8. Prepare a Peer Review Consensus Report summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. (Drafted by the Panel during the Review Workshop with a final report due three weeks after the workshop ends.)

A first draft was completed during the RW. All Panel members contributed sections and the Assessment Team provided text and plots related to requests during (and after) the RW (see Appendix A). The report was finalized by email after the RW.

3. Additional comments

Model runs which excluded age-0 snapper were included in the Advisory Report, after the RW, by the Assessment Team. The Panel wished to emphasize their preference for the inclusion of age-0 snapper in the assessment and requested results for the RW's base case for inclusion in the Consensus Summary Report. These results are briefly discussed below.

The benchmarks are dependent on the assumed effort allocation and this must be kept in mind when considering the RW's base case results (Table 1). Note, the benchmarks have been calculated, and projections have been done, assuming that recruitment is equal to the average of the base case estimates from 1984–2003. Under the linked effort scenario, the SPR at F_{MSY} is 27% and hence the 30%-SPR results are very similar to results for runs where MSY was the benchmark (Table 1). For the current shrimp scenario and the 40% shrimp effort reduction, the rebuild to MSY levels occurs much more quickly than for the linked scenario (Table 1). However, this is because the target levels are much lower (5-10% SPR compared to 27% SPR, see Table 1). Note, SPR values of 30% or higher can only be achieved under the linked scenario.

The RW noted that although there were periods of time when there was good sampling coverage from a range of sources and fisheries, this was generally not the rule for red snapper (e.g., patchy age data with missing years). The consistent and sustained collection of data for stock assessment purposes is a generic issue for GoM species. Good quality data over an extended timeframe needs to be available for monitoring purposes and for stock assessment as the need arises. We understand that work is currently underway to address sampling needs for a range of species. The RW supports such statistical studies to provide sampling specifications to data providers, so that sufficient age- and length-composition data are available for assessments. It may also be timely to review protocols for ensuring random (representative) sampling in the various fisheries and monitoring programs.

The Panel made several requests of the Assessment Team for additional analyses, including some additional model runs. The requests are listed below. Further details, the results, and discussion of the results are given in Appendix A.

<u>Model runs</u>

1. Initial base case with high virgin recruitment and constrained directed catch history

2. ASAP run with revised and expanded input data (1984 time series) and revised parameters but the same approach as the 1999 assessment

3. All remaining runs for combinations of:

{1872 time series, 1984 time series} $\times$ *{const q, random q}* $\times$ *{age 0, age 1}*

The shorthand "random q" refers to allowing a random walk in the catchability coefficients of the directed fisheries; "const q" denotes constant catchability in the directed fisheries. For x = 0 or 1, "age x" denotes age-x red snapper as the minimum age class in the model. Three of these runs had already been completed. The remaining runs were prioritized with the two runs "1984 time series – random q – { age 0, age 1}" given the lowest priority.

4. Projections from the base case using higher average future recruitment

The Panel requested additional projections for the base case where future recruitment and MSY calculations were predicated on higher values for R_0 than the estimated value. The requested alternative values for R_0 were (a) the average of the estimates from 1984 to 2003 and (b) the value estimated with the 1984 time series. Two sets of deterministic projections, based on recruitment scenario (a), were completed in time to be shown to the group prior to the close of the RW meeting.

<u>Diagnostics</u>

1. Standardized residuals: Q-Q plots and standard deviation of standardized residuals

2. Capture rate (catch + discards) at-age trajectories

3. Spawner-recruitment relationship

<u>Miscellaneous requests</u>

1. Mature biomass trajectories in contrast to effective-spawner trajectories

2. Virgin predicted selected age frequencies (by fishery) contrasted with average observed age frequency

There were two minor analyses which were undertaken by Panel members.

1. The absence of a plus-group, at the maximum age of 30 years, in the population model was of concern because of the relatively low assumed adult natural mortality (0.1). It was possible that the cumulative number of fish aged 31 years or older might be sufficiently large to unduly bias estimates of ratios involving the virgin stock. However, when this was checked, for effective spawners, the bias was found to be only 10%, which is inconsequential for the current assessment.

2. The figures and tables in workshop documents presenting catch-at-age estimates were not adequate for the purpose of checking, by eye, for the presence of consistently strong or weak cohorts. To alleviate the workload of the Assessment Team, a Panel member produced bubble plots of age vs cohort and presented them to the RW. By eye, it was difficult to detect any consistently strong or weak cohorts. The presentation was ideal, and illustrated the strongly domed selectivity pattern in the main fisheries. However, a domed selection pattern reduces the number of times that a cohort is seen in a fishery and this obscures consistent strength or weakness. That said, it was not obvious that the observed data were entirely consistent with the highly variable pattern of recruitment estimates in the assessment runs. There is a case for further investigation of residual patterns for the catch-at-age data to check, amongst other things, that recruitment estimates are being driven by appropriate time series (i.e., not by random fluctuations in abundance times series).

4. Stakeholder comments

As an industry representative of the for-hire fishery to the Review Panel and chairman of the GMFMC's Red Snapper Advisory Panel, here are my non-scientific feelings about the meeting.

The most disturbing thing that I encountered was the fact that the AW had decided to use a model that failed to include age-0 red snapper. The rationale was that due to high natural mortality rates of age-0 fish it was best, from a modeling standpoint, to begin the process with age-1 red snapper. The problem that I, and other non-scientists, have with this approach is that for the past twenty years we have been told that due to the high (80%) rate of shrimp by-catch mortality inflicted on the age-0 red snapper, rebuilding the stock could never be accomplished without very significant shrimp trawl by-catch reduction, regardless of what the directed fishery did or did not do.

Because of this, I am totally opposed to an assessment being released that omits the age-0 fish. The reason being that no matter how little effect age-0 omission would have on the actual model, it has the potential to have a huge effect on user group allocations when the Council begins using assessment to manage the red snapper stock.

Another thing that puzzles me as a non-modeler is the steepness of the recruitment curve. In an effort to understand and/or deal with or modify this recruitment steepness, there have been numerous runs and re-runs with different things factored in. The one thing that was never brought up is that maybe the stock has more spawners in it than are being accounted for and hence, at least in my mind, maybe the stock is in better condition then the model is showing. One of the Panel observed that it was hard for him to acknowledge the presence of an overfished stock with the recruitment steepness being shown by the model.

Another thing that I would like to speak about is the shrimp effort and/or by-catch reduction. While it is evident that BRD reduction rates are much lower then was hoped for and predicted, it seems to me that we must, somehow, find a way to incorporate the massive reductions (25%) of effort because of the economic upheavals in the shrimp fleet. These factors being high fuel costs, low shrimp prices, low performances of BRDs and the market glut of foreign and pond-raised shrimp. It seems to me that with the myriad of things that can be formulated and injected into the model, that this effort reduction can be computed as well.

Another concept that I feel is worthy of consideration is that possibly a reduction or elimination of minimum lengths might provide enough benefits in bycatch reduction and therefore by-catch mortality, particularly in the recreational sector, to offset the increased harvest that might result from such an action.

Mike Nugent, Chairman
Red Snapper Advisory Panel

5. Recommendations for future workshops

The RW has two major and two minor recommendations for future SEDARs. The recommendations are listed below followed by their justification.

1. *Change the Review Panel instructions to specifically allow minor changes to the assessment in collaboration with the Assessment Team.*

During the RW the Panel identified what were, in their opinion, deficiencies in the assessment. A strict interpretation of the Panel's instructions would have required that the shortcomings be noted in the Consensus Summary Report together with suggested remedial actions. According to their instructions the Panel was not able to request an alternative assessment. However, the remedial actions were minor in nature, and the Assessment Team were willing to make the changes during the RW. The alternative of reconvening the AW and the RW in the future would have been inefficient in terms of time and money.

The RW acknowledges that by opening the door to "minor changes" that a grey area is introduced. However, the Panel believe that future Review Panels should be attributed with sufficient common sense to allow them some latitude. They should always be guided by whether changes to the assessment can be made "safely" (without an undue possibility of errors being made), are in the spirit of the assessment (i.e., not using a different method or model), and are agreed to by the Assessment Team and the SEDAR Coordinator.

2. Review RW Term of Reference 5 to bring the Advisory Report back into the RW Terms of Reference.

The RW Term of Reference 5 has two parts. First, there is a check that the "Stock Assessment Report" is consistent with the required outline. Second, there is a check that the results are consistent with the Panel's decisions regarding adequacy and appropriateness. The second part creates some problems if an RW finds any deficiencies with the assessment which are addressed during the RW. Should the Assessment Report be revised to include the new results? This Term of Reference makes more sense if an RW is not a workshop, but simply an "accept" or "reject" forum.

The assessment goes forward into the Advisory Report, but an RW does not consider the Advisory Report in any of its Terms of Reference. During this RW, the base case and sensitivity runs to be presented in the Advisory Report were recommended by the Panel and agreed to by the Assessment Team (although the Assessment Team also included runs in the Advisory Report that the AW had agreed upon). The RW spent some time reviewing the Advisory Report (although the full set of results were not available). We took this approach because it seemed appropriate that an RW's decisions are necessarily reflected in the Advisory Report. That is, the possibility of a disjunction between an RW's decisions and the Advisory Report should be minimized.

3. Clarify Data Workshop and Assessment Workshop Terms of Reference

The RW had some difficulty in understanding the exact purpose and meaning of some of the DW and AW terms of reference. We suggest a brief review of these terms of reference.

4. Send documents as electronic copies, with hard copies of the main reports only.

This would provide some cost savings without detracting from the information available to participants. If a participant really does require all documents in hardcopy, they could still be provided on request.

Appendix A: Summary of Assessment Team results in response to Panel requests

Model runs

1. Initial base case with high virgin recruitment and constrained directed catch history

The Panel wanted some confirmation of why the input data necessarily lead to high current depletion. In an attempt to clarify this issue we requested a model run which fixed virgin recruitment at a much higher level than was estimated in the initial base case. The CVs on the directed catch history were modified to force the directed catches to be taken. The expectation, of some Panel members, was that there would be a bad fit to some, or most, of the abundance indices (showing that the indices were incompatible with a much larger virgin stock size).

The model found a best fit to the data by estimating a long series of poor recruitment from the beginning of the fishery (1872) up to near the beginning of the available abundance data. Predicted shrimp bycatch was reduced but the fit to other data was similar to the initial base case. The Panel did not pursue further runs aimed at understanding why the data were producing the high current depletion. We concluded, that with flat or increasing abundance indices in recent times, that fishing down had to have occurred before the period of the abundance indices, and that recent high catches were necessarily supported by good recruitment. The level of depletion was probably dictated by the extent of truncation in the catch-at-age data.

2. ASAP run with revised and expanded input data (1984 time series) and revised parameters but the same approach as the 1999 assessment

The Panel wanted to understand what the primary differences were between the previous assessment results in 1999, and the current assessment results, and whether the differences were due to a change of model or data. A single extra run was proposed, termed the “continuity run”. This was specified to incorporate all input data (1984 time series) and parameter changes adopted in the current assessment, but to use the model (ASAP) and “logic” of the 1999 assessment.

There are a number of difficulties when making comparisons with the 1999 assessment results. First, there were “low” and “high” recruitment scenarios considered in 1999. For the continuity run, the same logic was applied, as in 1999, to derive low and high recruitment runs for comparison. However, the logic of 1999 delivered different values of R_0, than those obtained in 1999, for “low” and “high” recruitment. Secondly, ASAP and CATCHEM have different definitions of effective spawners. Comparisons between ratios are appropriate, but absolute values cannot be compared. For this reason, absolute comparisons were made using mature biomass. Lastly, the CATCHEM base case has eastern and western stocks, but in 1999 there was a single stock assumption. Comparisons are made, where appropriate, by summing eastern and western estimates.

The continuity run with high recruitment gave almost identical estimates of depletion to the 1999 high recruitment run (Figure 1). The low recruitment runs gave similar estimates of depletion, in an absolute sense, but showed different trends (Figure 1). When considered relative to an S_{MSY} benchmark the continuity runs are somewhat different to the 1999 runs, in an absolute sense, but show very similar trends (Figures 2 & 3). The CATCHEM base case shows less depletion than the ASAP runs, ranging from 1–8% of virgin effective spawners (Figure 4). However, all of the runs show high levels of depletion (less than 10% of virgin effective spawners, see Figures 1 &

4). In terms of mature biomass, large differences are seen between the 1999 ASAP runs, the continuity runs, and the CATCHEM base case (Figure 5). The CATCHEM run shows the lowest estimated levels (from 1989 onwards), with the 1999 ASAP runs being higher by a factor of 3–4 (Figure 5). About half of the difference is accounted for by the change in the maturity and mean weight-at-age vectors (see Figure 6).

3. All remaining runs for combinations of:

{1872 time series, 1984 time series} x {const q, random q} x {age 0, age 1}

The shorthand "random q" refers to allowing a random walk in the catchability coefficients of the directed fisheries; "const q" denotes constant catchability in the directed fisheries. For x = 0 or 1, "age x" denotes age-x red snapper as the minimum age class in the model. Three of these runs had already been completed. The remaining runs were prioritized with the two runs "1984 time series – random q – { age 0, age 1}" given the lowest priority.

The three dimensions of the eight runs were identified as the primary "dimensions of choice", and the RW agreed that one of these runs would be selected as a base case (such a selection was a milestone in the draft RW Agenda). The length of the time series (primarily catch history) is an important choice because it must be acknowledged that the early catch history, although based on best available data, has uncertainties associated with it which cannot adequately be captured by assigning relatively arbitrary (but high) CVs. The 1984 time series option uses only actual observations. The random walk q was investigated as there undoubtedly have been changes in catchability (due to technology improvements). There was concern that the higher recent recruitment estimates could be an artifact of the model assumption that restricted catchability to a constant level. The issue of age-0 fish being included or not is clearly important (see Section 2.2).

The two random q runs with the long catch history were found to be very similar to the constant q runs. While these sensitivities suggested that catchability may have been changing, any conclusions are weak because of the lack of direct observations on fishing effort. The sensitivity runs made no substantial difference in the estimated recruitment pattern. Because of these results, the request for the two low priority runs was withdrawn

The length of the time series made some difference to the absolute level of biomass (and hence long term yields) but gave similar results with regard to depletion level. The previously observed instability of the solution to the 1984 time series was still present, and the likelihood surface was perceived as being much "flatter". The omission or not of the age-0 red snapper in the model made little difference to a qualitative assessment of the results. The RW chose as a base case the 1872 time series, with constant q, and inclusion of the age-0 red snapper. The 1984 time series, with constant q, and inclusion of age-0 red snapper was recommended as a sensitivity run to be taken forward to the Advisory Report.

4. Projections from the base case using higher average future recruitment

The Panel requested additional projections for the base case where future recruitment and MSY calculations were predicated on higher values for R_0 than the estimated value. The requested alternative values for R_0 were (a) the average of the estimates from 1984 to 2003, to reflect the possibility that the more recent values may provide a better reflection of recruitment in the near future, and (b) the value estimated with the 1984 time series, ostensibly as an upper bound. These requests required non-trivial changes to the existing code, which were accomplished towards the end of the meeting. Two sets of deterministic projections, based on recruitment scenario (a), were completed in time to be presented to the meeting.

The first set assumed a 40% reduction in shrimp bycatch rates beginning in 2007 and various levels of constant catch from the directed fishery. The results indicated that the stock could recover to MSY levels by as early as 2017 even with the current TAC, provided shrimp bycatch is in fact reduced by 40% (and provided post-settlement compensatory mortality effects are unimportant relative to shrimp bycatch). The second set of projections assumed current shrimp bycatch rates would continue into the future and the effort of the directed fisheries would be reduced to F_{MSY}. Under those conditions the stock could recover to MSY levels by 2025, but the initial TAC would have to be reduced to about 7 million pounds.

Diagnostics

1. Standardized residuals

The original assessment was rather weak on the provision of diagnostics. The Panel was interested in whether the residuals were consistent with the model's assumed (and estimated, through a common variance term) CVs and the statistical error structures: lognormal for catch, effort, and abundance indices; and multinomial for catch-at-age.

The production of quantile to quantile (Q-Q) plots was requested for the RW base case together with the standard deviations of the standardized residuals (sdsr). If the assumptions of the model are satisfied then Q-Q plots should show the residual distribution near the y=x line, and the sdsr values should be near to 1.

Most Q-Q plots showed good agreement with the lognormal assumption. The multinomial assumption for the catch-at-age data did not appear to be satisfied (Figure 7). The distribution of catch-at-age residuals was skewed with a standard deviation much greater than 1 (Table 2). Most other time series had residuals consistent with their CVs, the exceptions being the two handline time series (which were fitted too well relative to their CVs) and the larval-E time series (which was fitted badly relative to the CVs).

The Panel did not consider these results to be a problem for this assessment. Rather, they viewed the further development of diagnostics as work for the future.

2. Capture rate (catch + discards) at-age trajectories

The Panel debated what would be a useful diagnostic for a reality check on the estimated catch levels. The question is whether estimated catch levels are credible given the available biomass. The Panel requested time trajectories of (instantaneous) capture rate (catch plus discards) at age by stock.

The two stocks showed different patterns at age as would be expected given different levels of shrimp bycatch and the somewhat different selectivity patterns of the fisheries (Figure 8). The eastern stock had lowest rates on ages 0–2 years, with highest rates on ages 3–5 years (Figure 8a). In contrast, the western stock had its highest rate on age-1, with lowest rates on the oldest age classes; the age-0 red snapper had rates similar to ages 3-7 years (Figure 8b). The credibility of any of these rates was not addressed by the RW as there is currently insufficient understanding of the distribution of age classes relative to the effort in the fisheries.

3. Spawner-recruitment relationship

The Panel requested, for this report, a plot of the estimated recruitment used in the RW base case, together with the predicted average future recruitment from the S-R function (Figure 9). This plot illustrates, for both stocks, that the past and future recruitment from the S-R function is lower than average estimated recruitment over the last 20 years. This is why, on the balance of probabilities, the Panel prefer the use of mean estimated recent recruitment to predict future recruitment levels.

Miscellaneous requests

1. Mature biomass trajectories in contrast to effective-spawner trajectories

The Panel wanted some idea of the effect of increasing egg-production at age on the perception of stock depletion. That is, what proportion of the high level of depletion in the total egg production (as measured by effective age-30 spawners), was due to the loss of older, larger, fish (females), and what was due to depletion of mature biomass.

The Assessment Team produced plots of mature biomass trajectories (as a proportion of virgin) for each stock which were based on mean weight-at-age from catch data, which was only available up to age 15. The comparison of mature biomass with effective spawners showed a divergence between the trajectories for the western stock early in the time frame (1870–2003) which was not present for the eastern stock. The RW concluded that the stock difference may have been an artifact of the use of mean weights from catch data. A more appropriate method of calculating mature biomass was pursued after the RW.

2. Virgin predicted selected age frequencies (by fishery) contrasted with average observed age frequency

The Assessment Team suggested it would be useful to contrast age frequencies for selected biomass in the virgin population with the average observed age frequency. The Panel agreed that this could provide some insight into how the observed age frequencies were influencing the estimates of depletion.

There is strong contrast between the virgin and exploited age frequencies even for the fisheries with highly domed selectivity patterns (Figure 10).

Table 1. Summary of results for the eastern and western stocks for the RW base case (age 0 included, 1872-2003 time series, R_0 = average recruitment from 1984-2003) for F_{MSY} and $F_{30\%}$ under the current shrimp effort, a 40% reduction in shrimp effort, and the current effort proportions ("linked"). SPR values of 30% or higher could not be achieved for the current shrimp and 40% shrimp reduction scenarios.

		Effort allocation schedule		
Area	Benchmark statistic	Current shrimp	40% shrimp reduction	Linked
East	MSY (mp)	4.6	5.4	6.6
	F_{2003}/F_{MSY}	2.3	2.1	3.8
	S_{2003}/S_{MSY}	0.34	0.34	0.12
	S_{2010}/S_{MSY}	0.7	0.67	0.42
	year $S/S_{MSY} = 1$	2020	2020	2027
	SPR at F_{MSY}	10%	10%	27%
	Yield at $F_{30\%}$ (mp)			6.6
	$F_{2003}/F_{30\%}$			4.1
	$S_{2003}/S_{30\%}$			0.11
	$S_{2010}/S_{30\%}$			0.39
	year $S/S_{30\%} = 1$			2027
West	MSY (mp)	7.1	12.1	19.9
	F_{2003}/F_{MSY}	2.3	2.1	3.8
	S_{2003}/S_{MSY}	0.26	0.17	0.04
	S_{2010}/S_{MSY}	0.62	0.42	0.24
	year $S/S_{MSY} = 1$	2025	2027	2032
	SPR at F_{MSY}	5%	7%	27%
	Yield at $F_{30\%}$ (mp)			19.8
	$F_{2003}/F_{30\%}$			4.1
	$S_{2003}/S_{30\%}$			0.04
	$S_{2010}/S_{30\%}$			0.22
	year $S/S_{30\%} = 1$			2032

Table 2. Standard deviation of the standardized residuals for each index and for all catch-at-age residuals. There are east (E) and west (W) series for each index (HL=hand line; LARV=larval survey; REC=recreational; TRW0=trawl survey age-0; TRW1=trawl survey age-1; VID=video survey).

Index	**Standard deviation of standardized residuals**
HL-E	0.53
HL-W	0.54
LARV-E	1.98
LARV-W	1.51
REC-E	0.71
REC-W	0.84
TRW0-E	1.47
TRW0-W	1.32
TRW1-E	1.08
TRW1-W	1.08
VID-E	0.81
VID-W	0.80
Catch-at-age	3.48

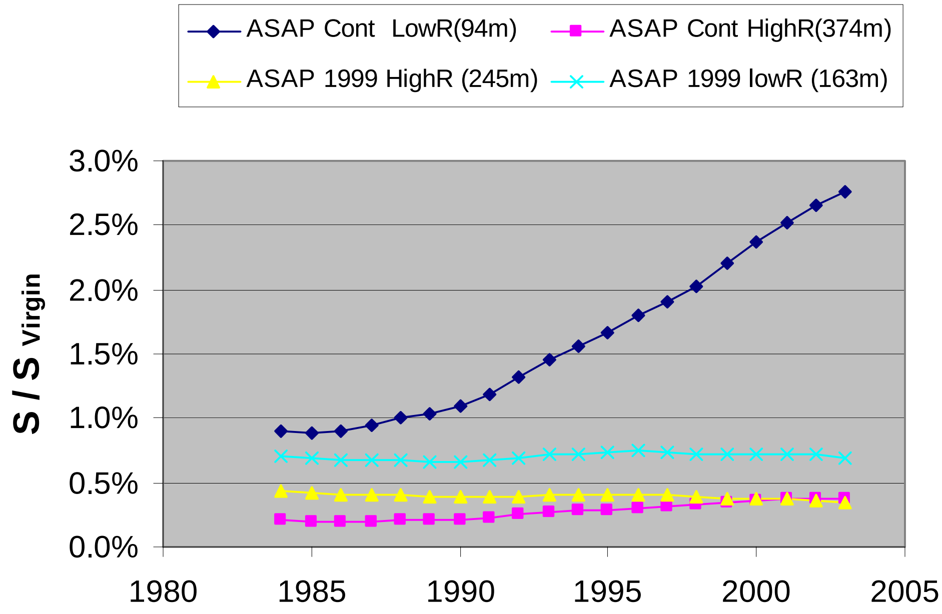


Figure 1: Effective spawners as a percentage of virgin from 1984–2003 for the 1999 ASAP assessment (ASAP 1999) and the RW continuity run (ASAP Cont). There are low and high recruitment scenarios for each case. Values of R_0 for the continuity run were derived using the same logic as the 1999 assessment.

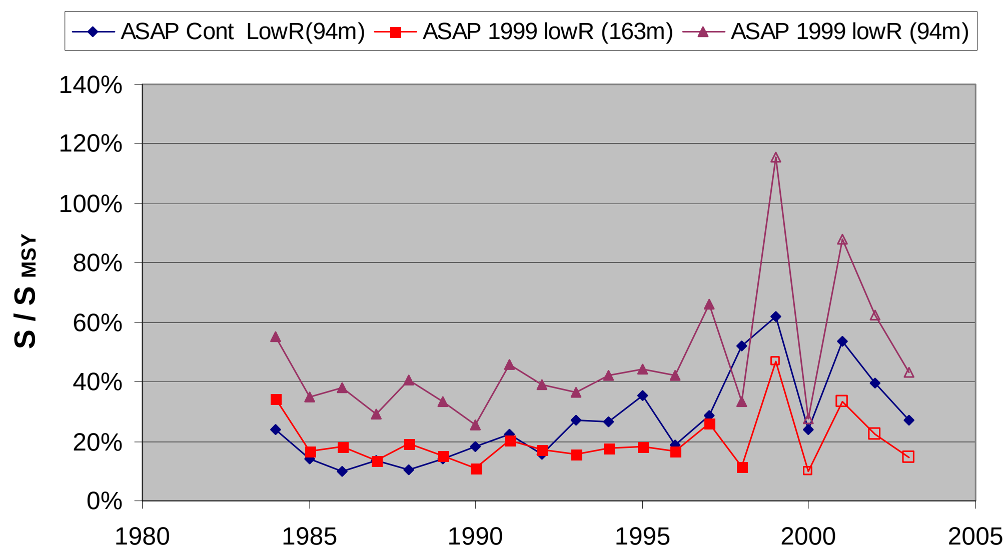


Figure 2: Effective spawners as a percentage of S_{MSY} for low recruitment cases. ASAP 1999 values for 1999–2003 are from a projection using observed directed yield and shrimp bycatch (indicated by open symbols).

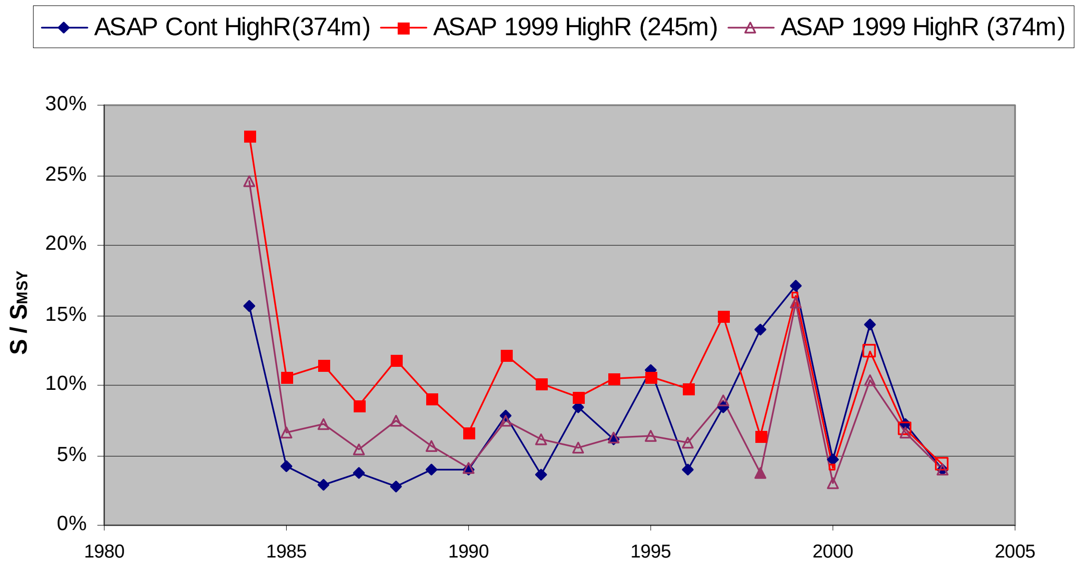


Figure 3: Effective spawners as a percentage of S_{MSY} for high recruitment cases. ASAP 1999 values for 1999–2003 are from a projection using observed directed yield and shrimp bycatch (indicated by open symbols).

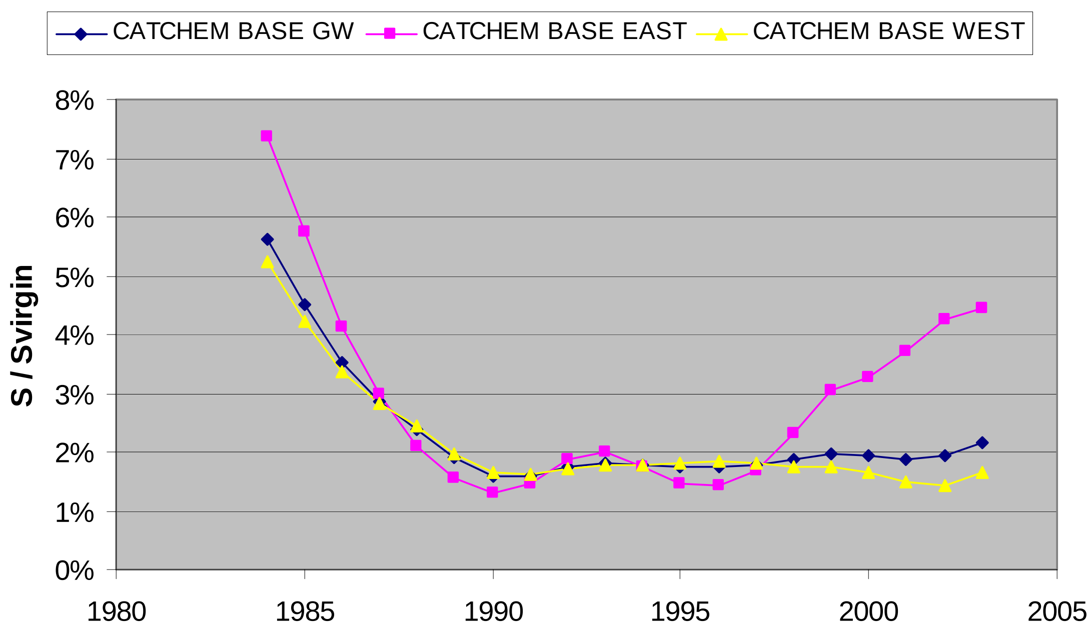


Figure 4: Effective spawners as a percentage of virgin from 1984–2003 for the CATCHEM RW base case. Results are shown for the eastern and western stocks separately and for the sum of the two stocks (GW).

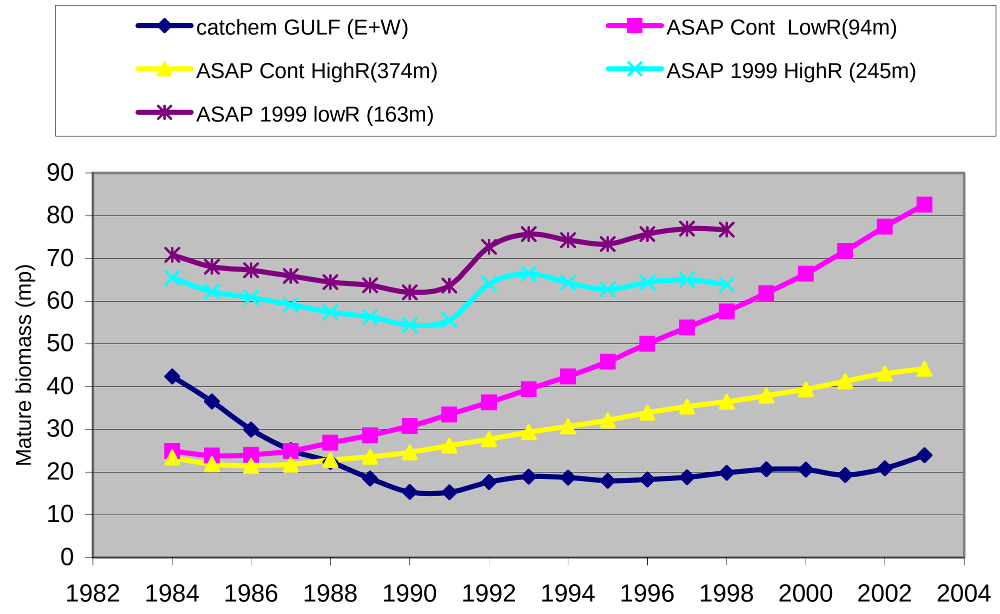


Figure 5: Mature biomass from 1984–2003 for the CATCHEM and ASAP continuity runs and for 1984–1998 for the 1999 ASAP low and high recruitment runs. Maturity and mean weight-at-age vectors used for the 1999 ASAP runs were consistent with assumptions in 1999.

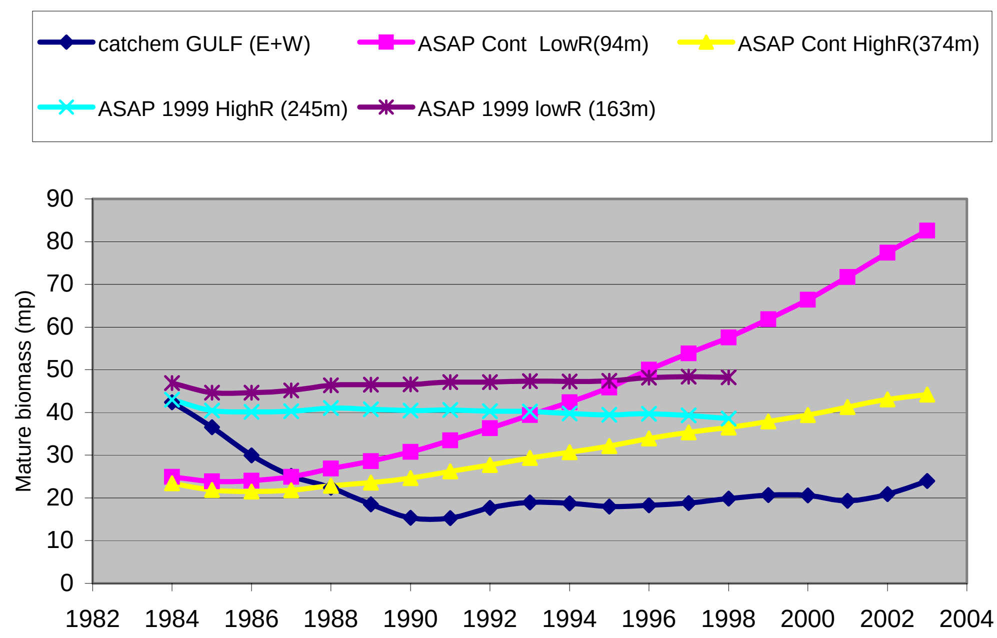


Figure 6: Mature biomass from 1984–2003 for the CATCHEM and ASAP continuity runs and from 1984–1998 for the 1999 ASAP low and high recruitment runs. The CATCHEM maturity and mean weight-at-age vectors were used for all runs.

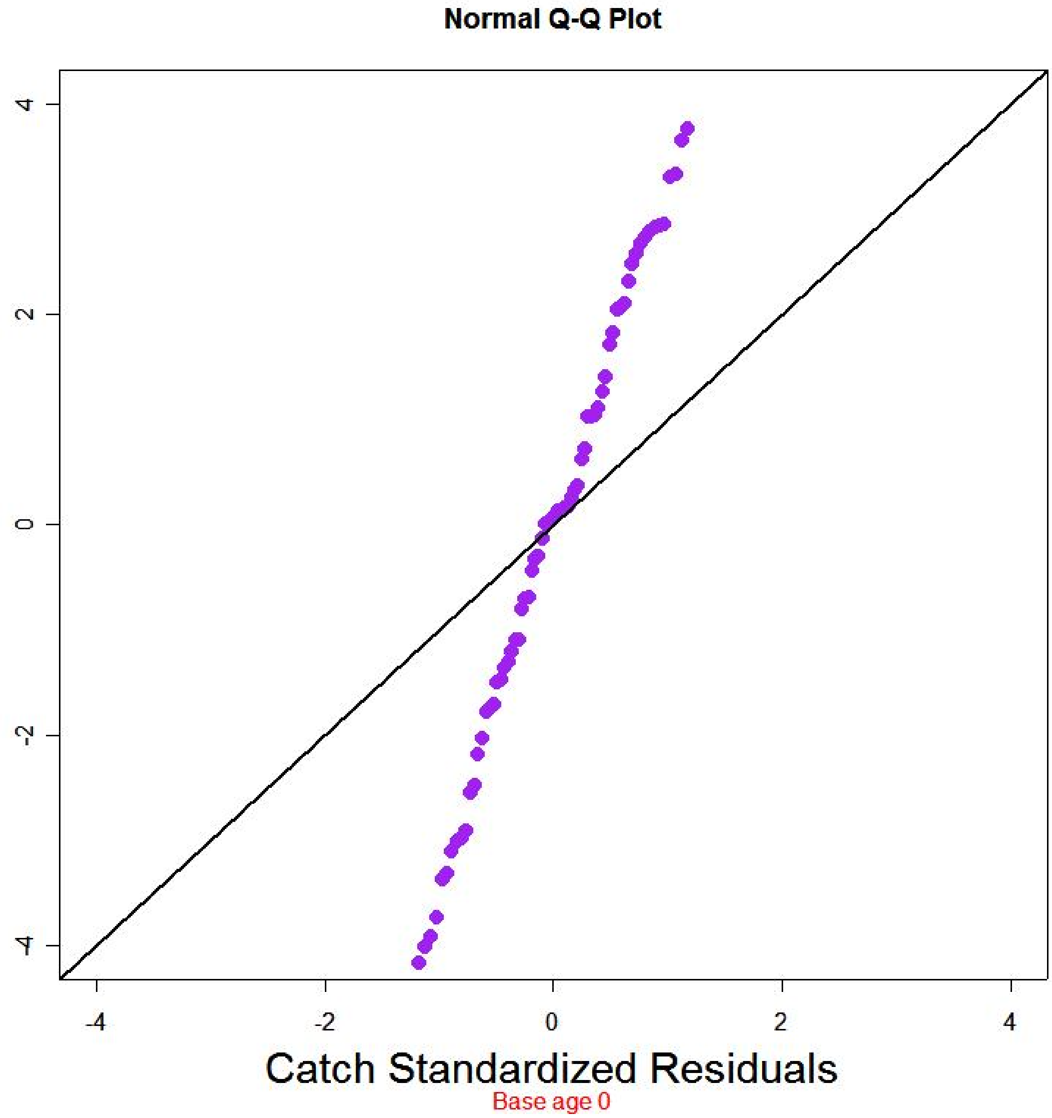
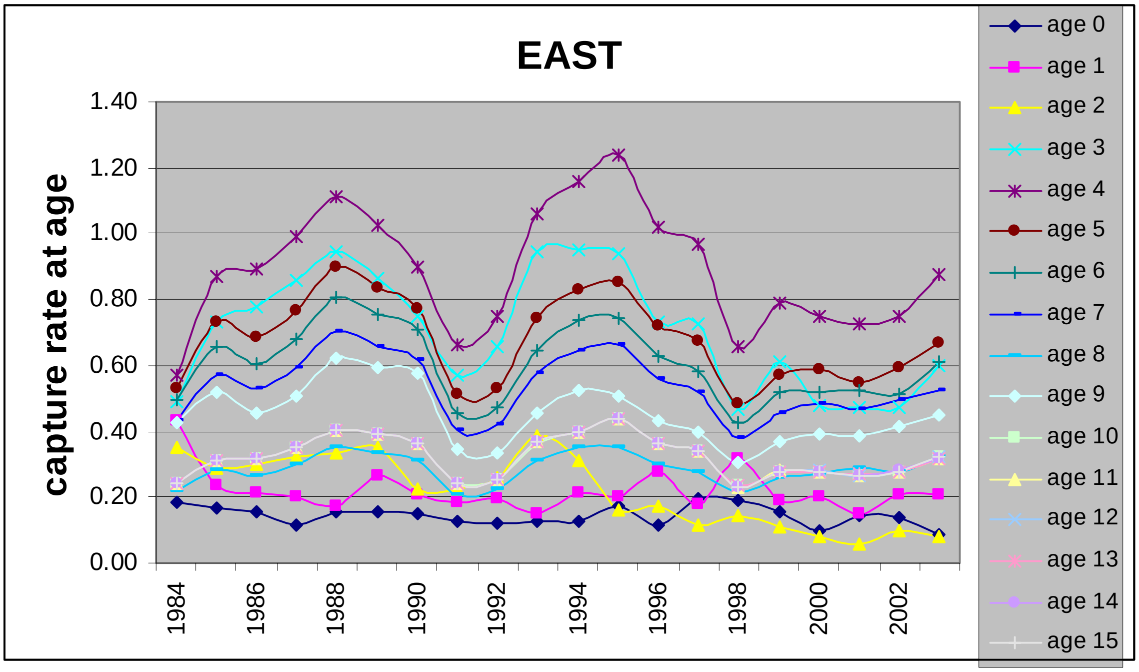


Figure 7. Q-Q plot of standardized residuals for the RW base case fit to observed catch-at-age (mean = -0.084, standard deviation = 3.48) showing highly skewed and over-dispersed residuals

(a)



(b)

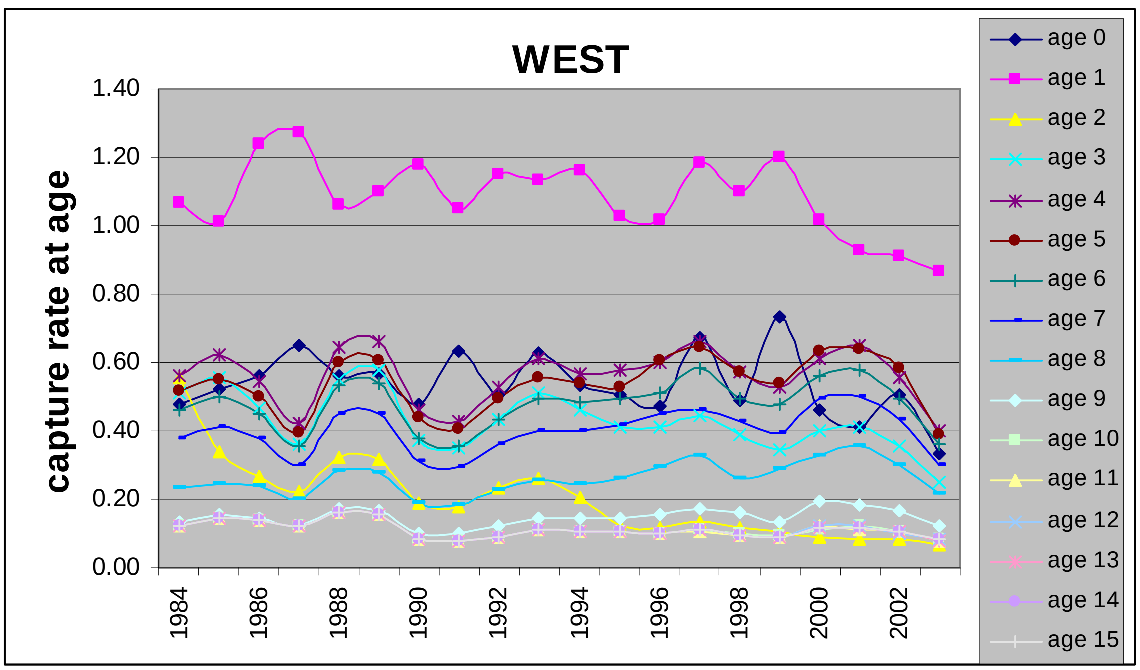
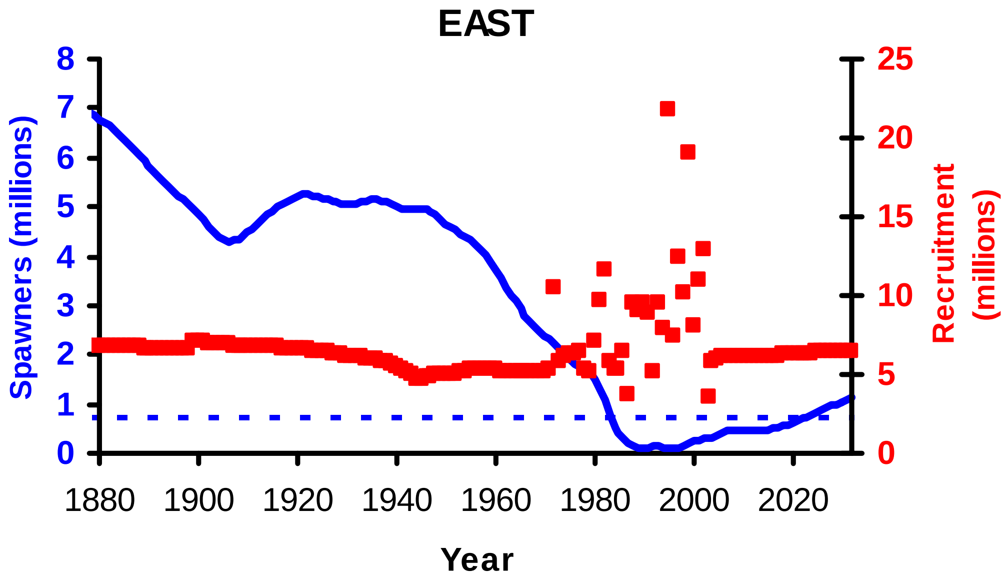


Figure 8. Capture rate at age in the east (a) and west (b) from 1984-2003. Capture rate reflects the instantaneous rate for fish that were caught (this includes landings as well as discards due to size limits and closed seasons). Age 15 is a plus group.

(a)



(b)

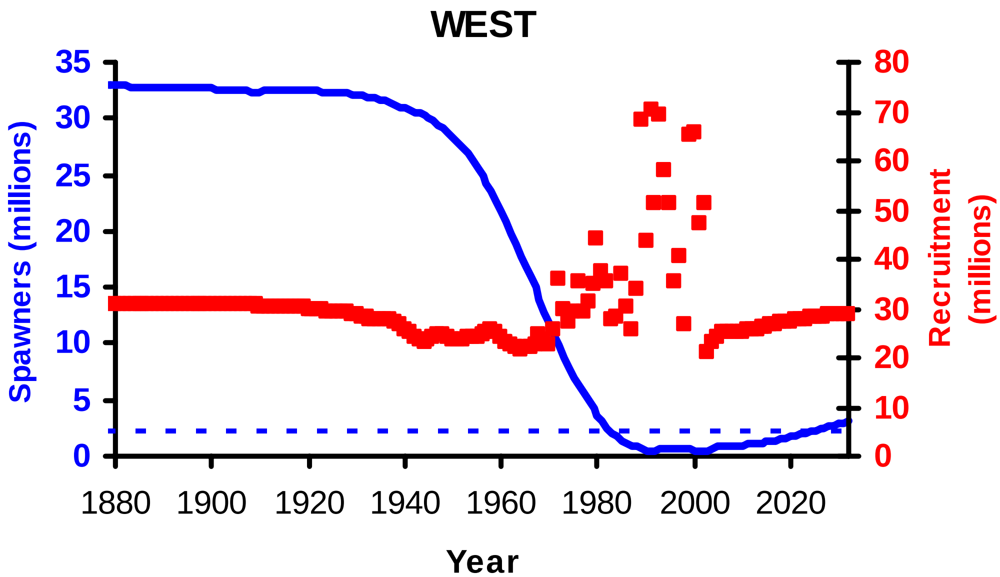
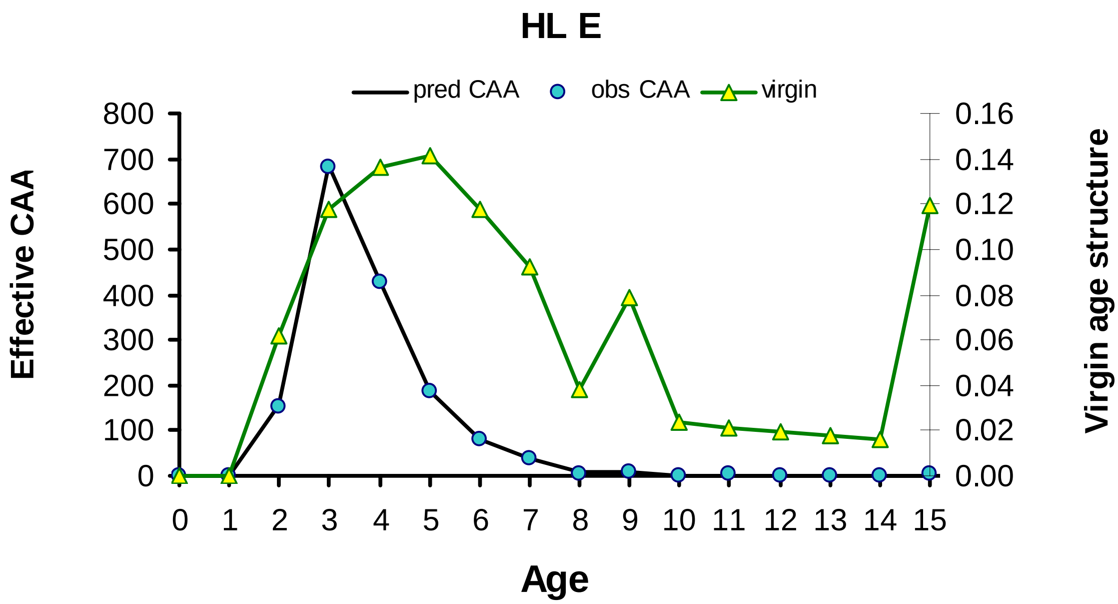


Figure 9. Trajectory of estimated effective spawners and predicted recruits in the east (a) and west (b) from 1872-2032. The dashed line is the effective spawners corresponding to 30%SPR.

(a)



(b)

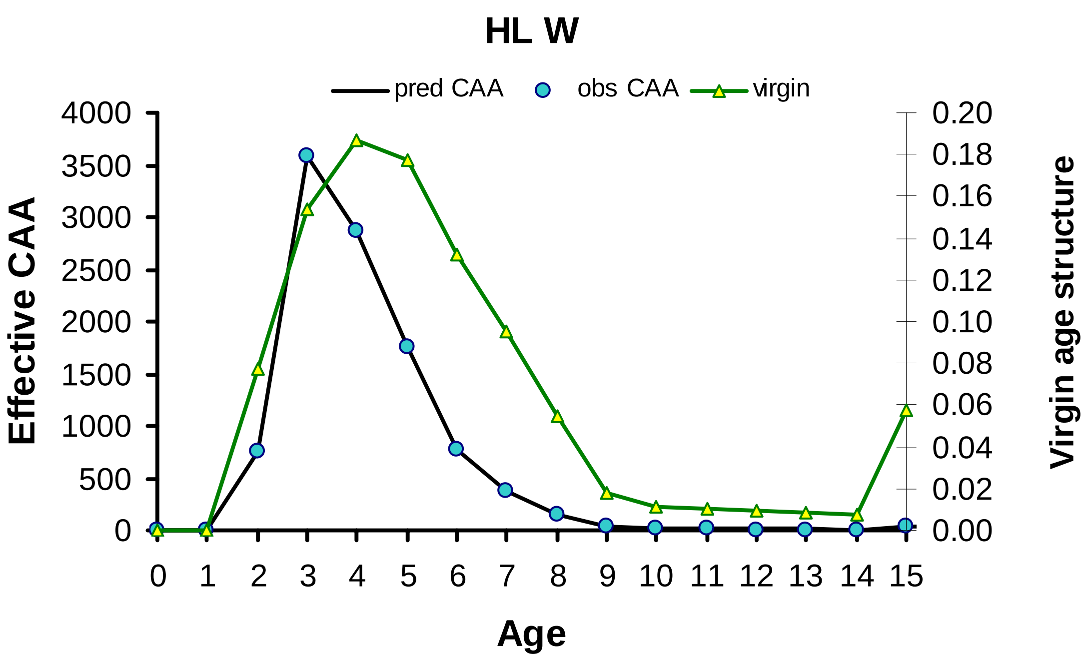


Figure 10. Unexploited age frequency (virgin) versus exploited age frequency (observed averaged across years, and predicted averaged across years) for the handline fisheries in the east (a) and west (b). Age 15 is a plus group.